Reservoir Computing Prediction of the Spread of COVID-19

Jacob Jeffries[1], Declan Norton[2], Tom Antonsen[2], Michelle Girvan[2], Brian Hunt[2], Andrew Pomerance[3]

[1]Clemson University [2]University of Maryland, College Park [3]Potomac Research

August 4, 2022

CLEMSON

TREND REU
Training and Research Experiences in Nonlinear Dynamics

The Problem

- We seek to accurately and efficiently forecast COVID-19
  - Policy-making, hospital preparation, vaccination efforts, etc.
  - Current ensemble model used by the CDC confidently forecasts a week into the future [Cramer et al., 2021]
- Can we predict a peak well before it happens, maybe weeks?
- Two primary challenges: **nonstationarity** and **bad data**

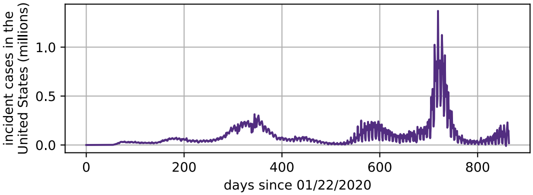

Figure: Raw national incident case data

CLEMSON TREND REU Training and Research Experiences in Nonlinear Dynamics

Background: Reservoir Computing

- Reservoir computers (RCs) are a type of artificial neural network
- **Goal**: Given a time series u_t, predict the next time step u_{t+1}

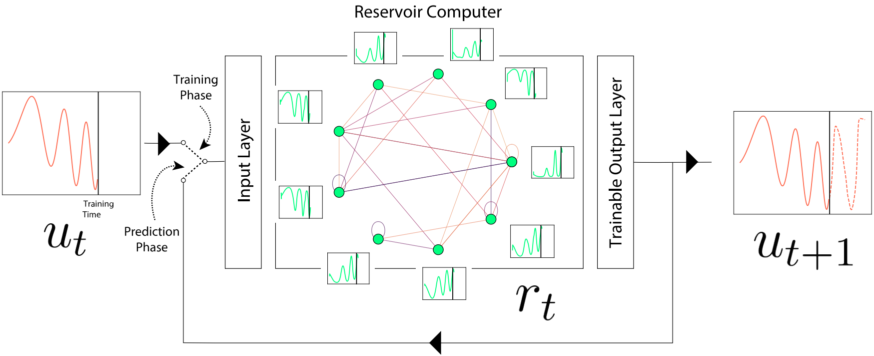

Figure: Reservoir computing schematic

CLEMSON TREND REU
Training and Research Experiences in Nonlinear Dynamics

- In the training phase, the output weights are set so that the next-step predictions best match the truth for the training data

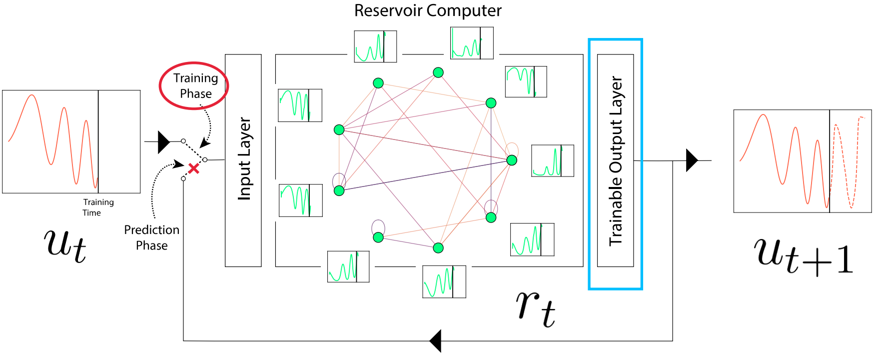

Figure: Reservoir computing schematic

- In the prediction phase, the output is fed back in as input to make longer term predictions

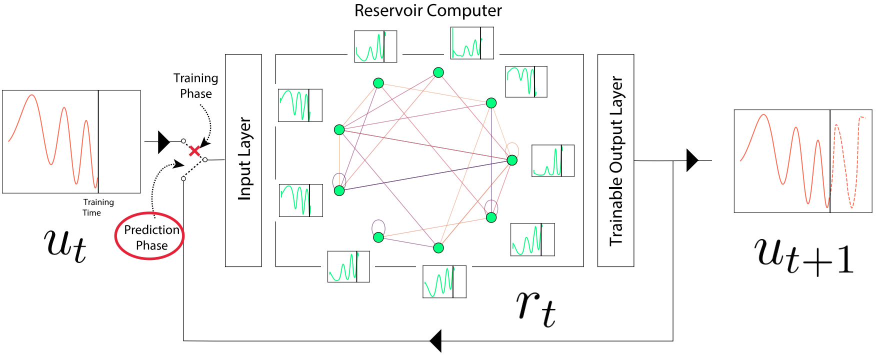

Figure: Reservoir computing schematic

Background: Reservoir Computing

Why reservoir computing?

- **Reservoir computing is efficient**
  - Training is only done on the output layer
  - This is in constrast to traditional deep learning, where multiple layers are trained
- **Reservoir computing is accurate**
  - Reservoir computing has shown impressive accuracy in predicting time series, accurately predicting some chaotic systems for multiple Lyapunov times [Pathak et al., 2018]

CLEMSON
TREND REU
Training and Research Experiences in Nonlinear Dynamics

Preprocessing

Raw data filtered by penalizing raw-filtered error and large fluctuations in the filtered form:

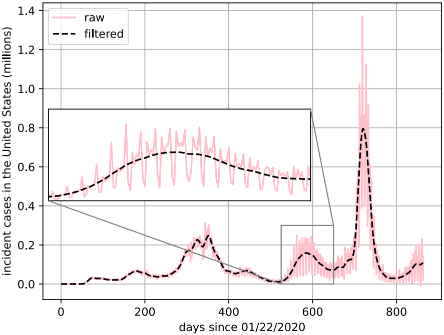

Figure: Filtered incident cases in the United States (x_t)

Preprocessing

- Seek to train a reservoir computer on "interesting" dynamics
- Identify time series surrounding peaks in **each state** (45 days in the past, 30 days in the future) found via a wavelet transform [Du et al., 2006]
- Detrend the data in each near-peak time series:

$$y_t = \log \frac{x_{t+1}}{x_t}$$

- where x_t is the filtered incident case time series
- Some reasons to detrend:
  - Each state has a different total population
  - Some waves are more extreme than others despite having similar dynamics

CLEMSON
TREND REU
Training and Research Experiences in Nonlinear Dynamics

- Train a reservoir computer on all near-peak detrended time series
- For comparison, we also performed the same procedure by training on all available data, segmented into 75-day windows

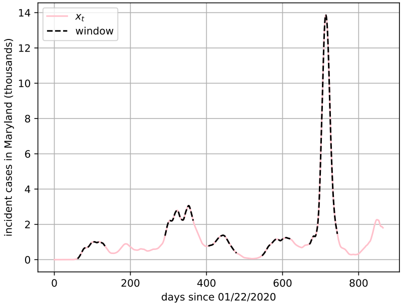

Figure: Each near-peak time series in Maryland

CLEMSON TREND REU Training and Research Experiences in Nonlinear Dynamics

Results

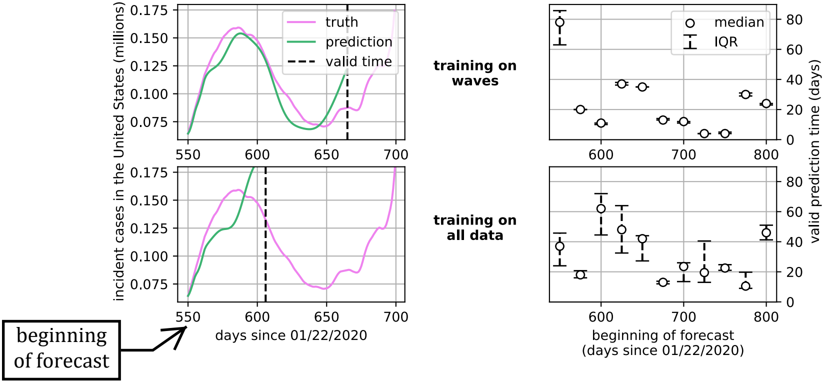

Figure: Example predictions in the United States, accurate for **weeks**

Future Work

- Couple the susceptible-infected-recovered (SIR) model with the reservoir computer to improve predictive strength [Pathak et al., 2018]

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -\frac{\beta S I}{N}$$
$$\frac{\mathrm{d}I}{\mathrm{d}t} = \frac{\beta S I}{N} - \gamma I$$
$$\frac{\mathrm{d}R}{\mathrm{d}t} = \gamma I$$

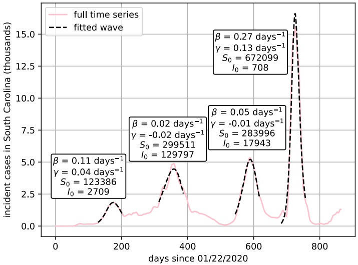

Figure: The SIR model (left) and fits to near-peak time series in South Carolina (right)

CLEMSON TREND REU Training and Research Experiences in Nonlinear Dynamics

Summary

- Reservoir computing accurately forecasts COVID-19 weeks into the future
- This beats the model ensemble used by the CDC and FiveThirtyEight, which has a confident prediction horizon of a week [Cramer et al., 2021]
- We expect even better predictive strength by coupling the reservoir computer with the SIR model

CLEMSON
TREND REU
Training and Research Experiences in Nonlinear Dynamics

Acknowledgements

- This work was sponsored by the National Science Foundation Award Number PHY2150399

- Special thanks to Declan Norton, Tom Antonsen, Michelle Girvan, Brian Hunt, and Andrew Pomerance

CLEMSON

TREND REU

Training and Research Experiences in Nonlinear Dynamics

References I

Cramer, E. Y., Huang, Y., et al. (2021).
The united states covid-19 forecast hub dataset.
medRxiv.

Du, P., Kibbe, W. A., et al. (2006).
Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching.
Bioinformatics.

Pathak, J., Wikner, A., et al. (2018).
Hybrid forecasting of chaotic processes: Using machine learning in conjunction with a knowledge-based model.
Chaos.

CLEMSON
TREND REU
Training and Research Experiences in Nonlinear Dynamics